cusp.ai

Machine Learning for Molecules and Materials

Max Welling

CuspAI

University of Amsterdam & Caltech

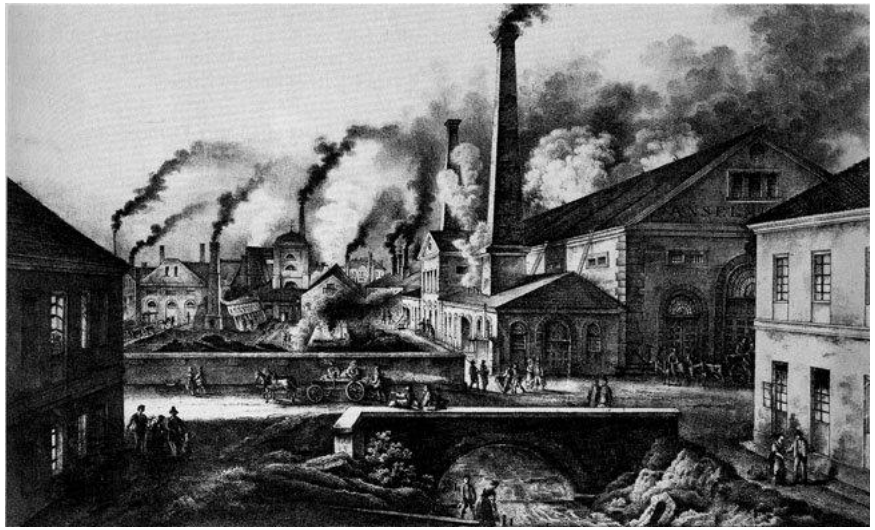
Future of Energy

March 10, 2025

Free Energy = Energy - Entropy

Ability to perform physical work

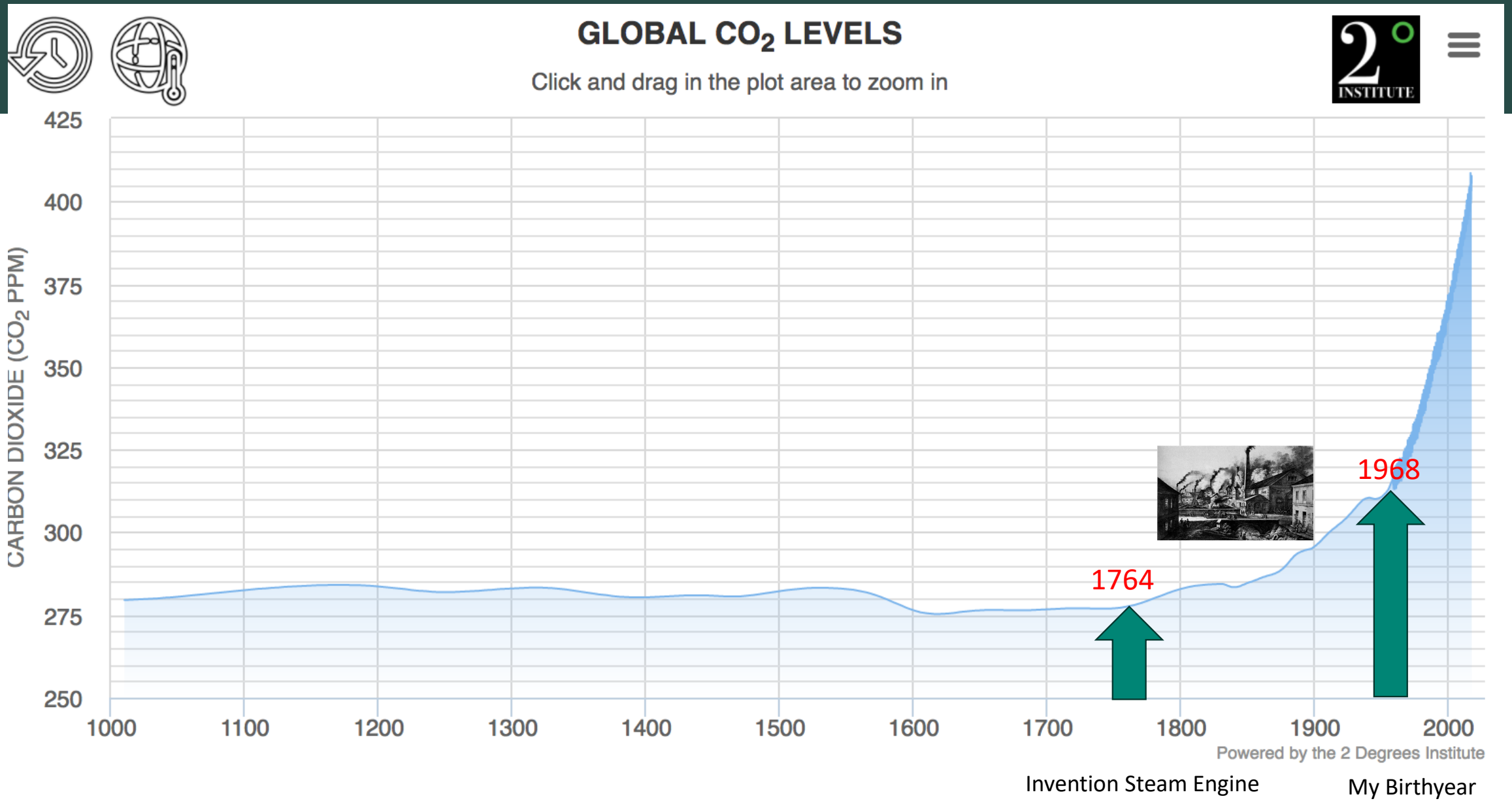
Level of organization, information of a system





Industrial Revolution: ±1820

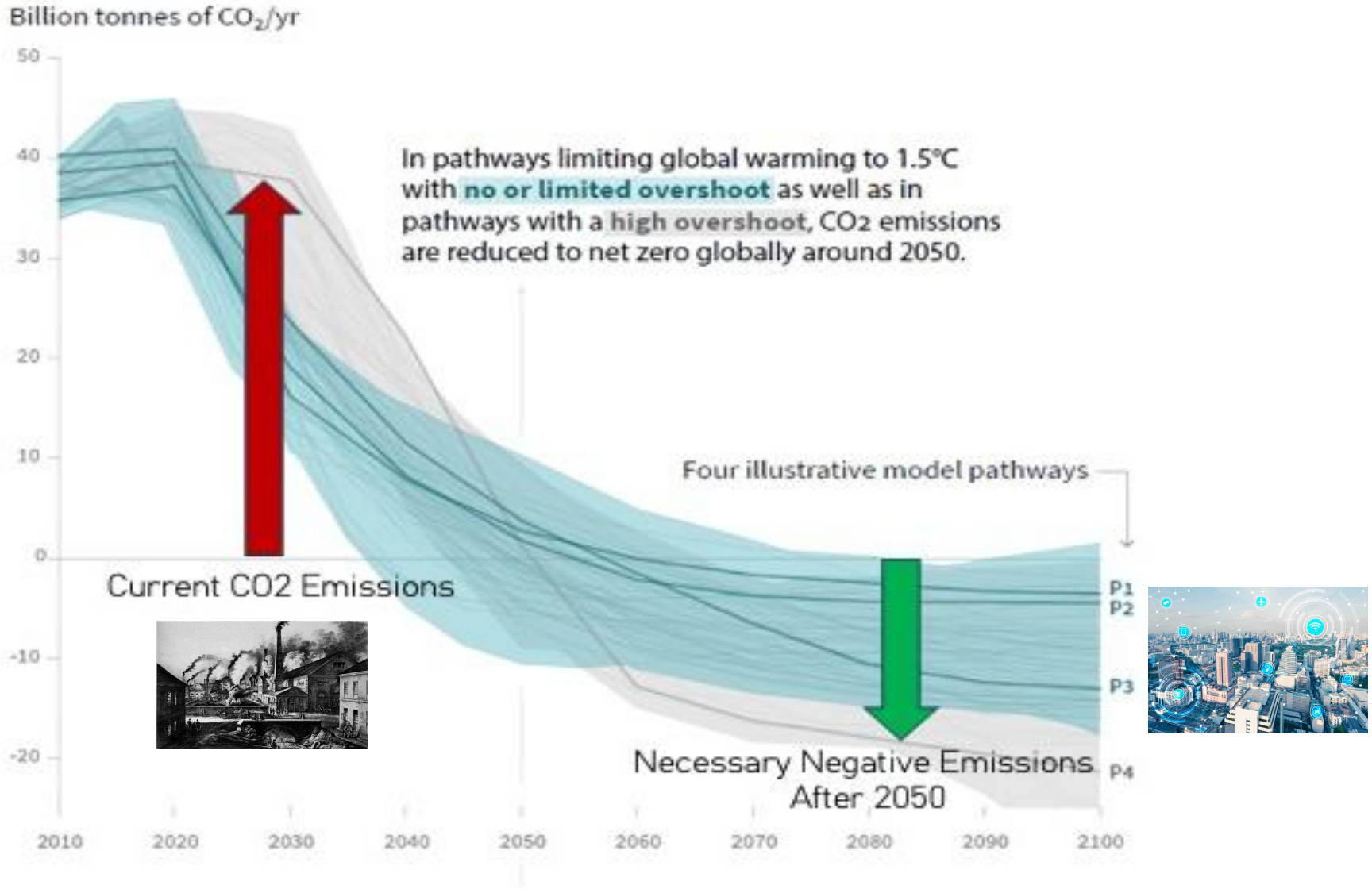
Information Revolution: ±1940



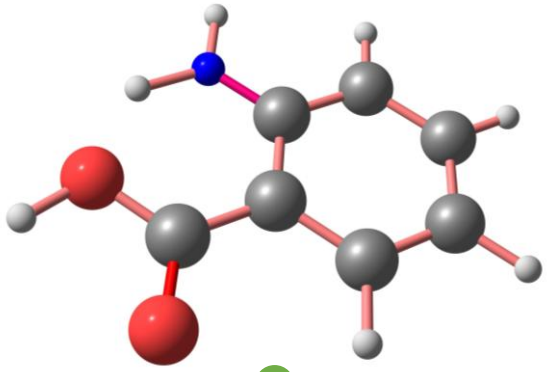
Invention Steam Engine

My Birthyear

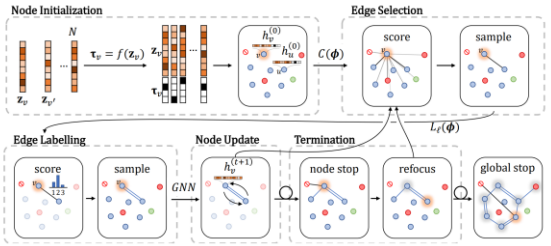
Necessary carbon removal to stay within 1.5 degrees

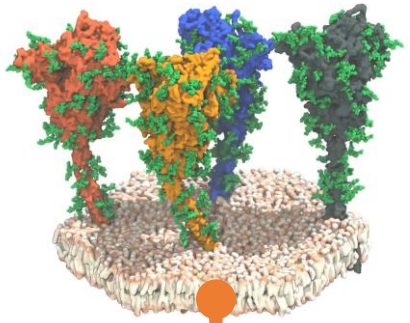


The Disruption in Two Steps:
Step I: From Experiment to Computer Simulation



data





Sören von Bülow, Mateusz Sikora, Gerhard Hummer.
MPI of Biophysics

Era 1: Trial-and-error

Era 2: Data-driven modelling

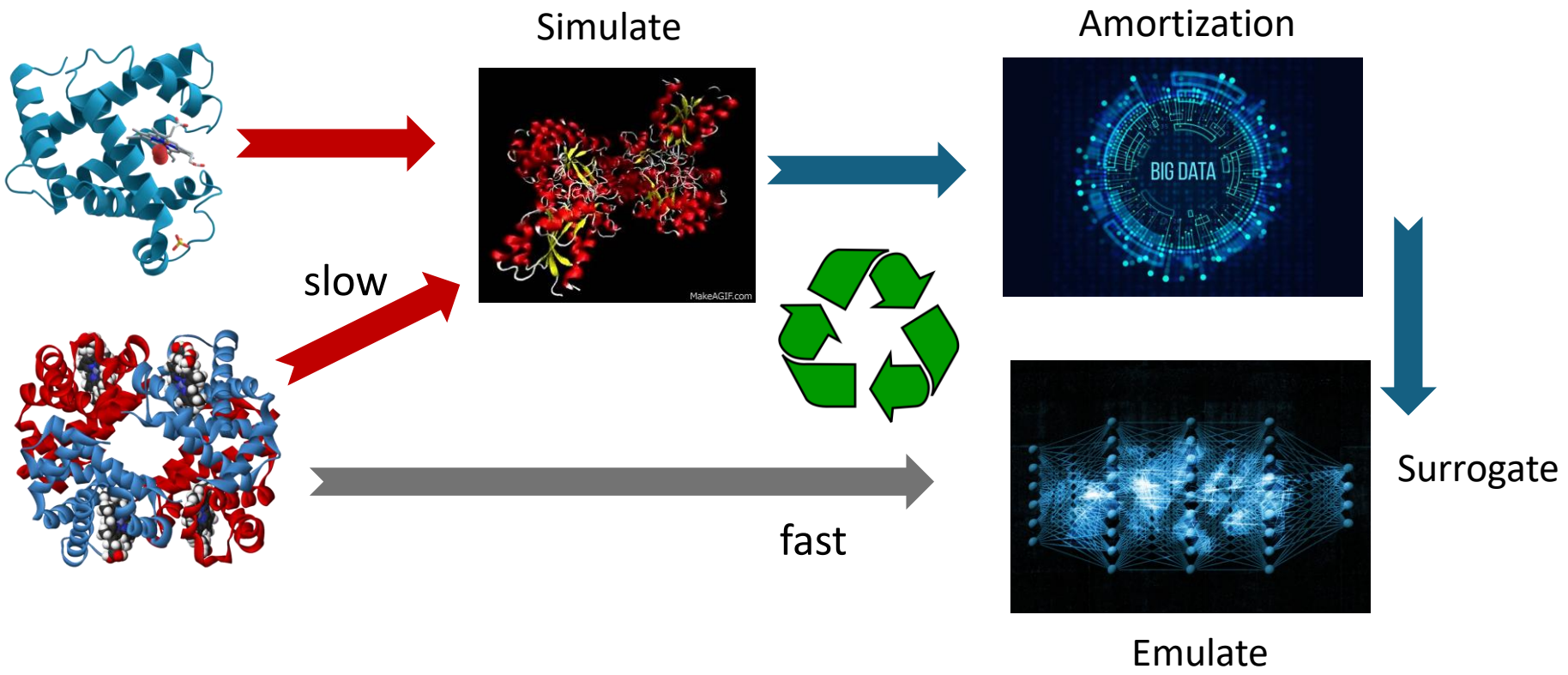
Era 3: In-silico design

COMPUTATIONAL COMPLEXITY

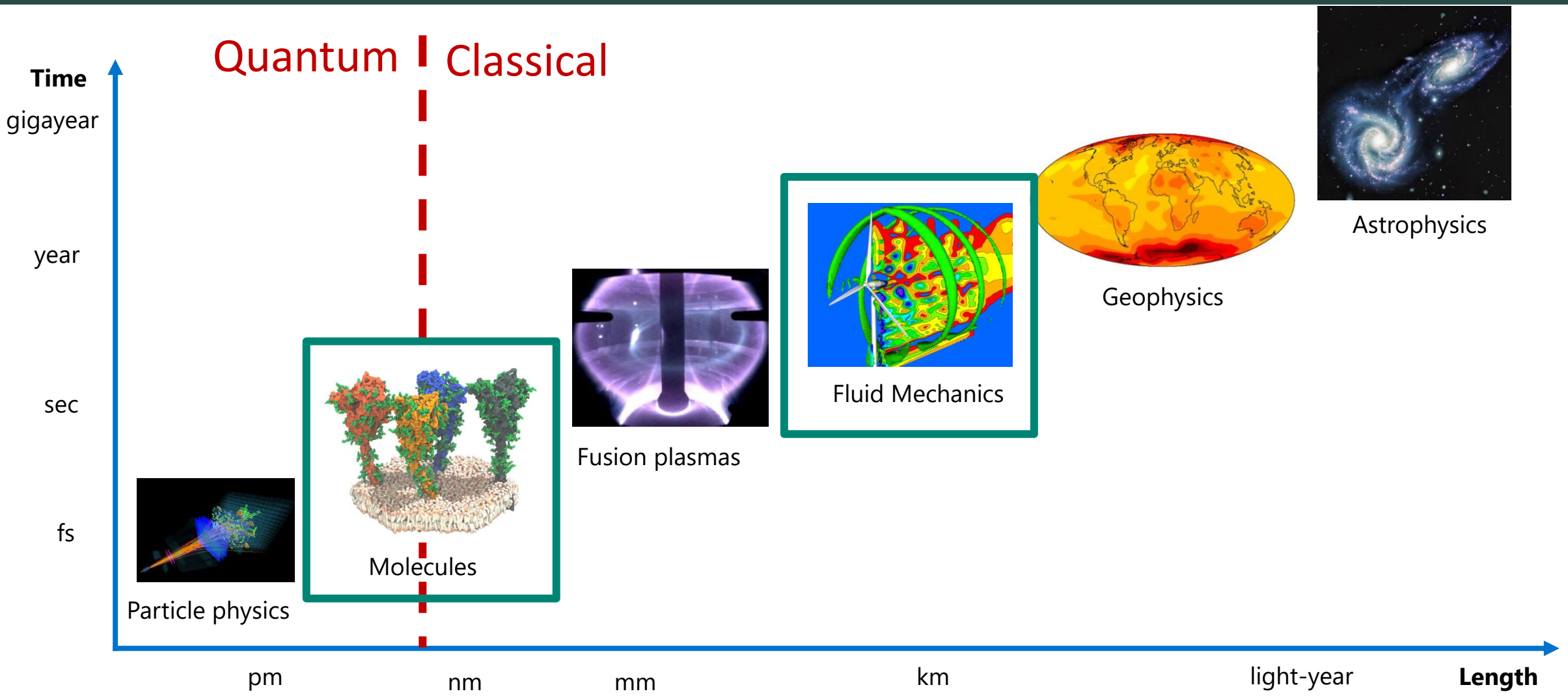
TIME

The Disruption in Two Steps:
Step II: From Simulation to Emulation

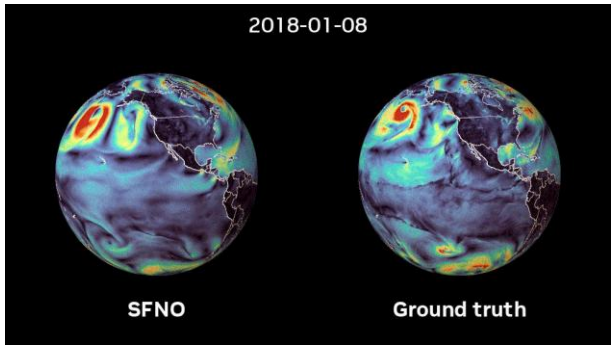
Simulate → Train NN Surrogate → Emulate

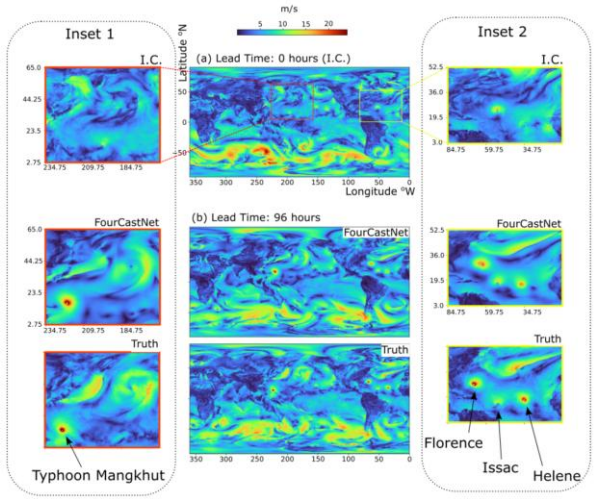


AI4Science: A Multi-Scale Scientific World



Weather Prediction with Neural Networks

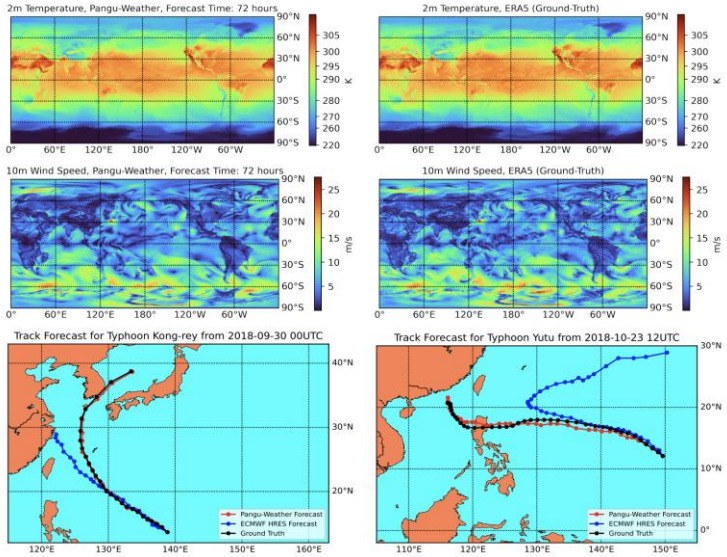




FOURCASTNET: A GLOBAL DATA-DRIVEN HIGH-RESOLUTION WEATHER MODEL USING ADAPTIVE FOURIER NEURAL OPERATORS

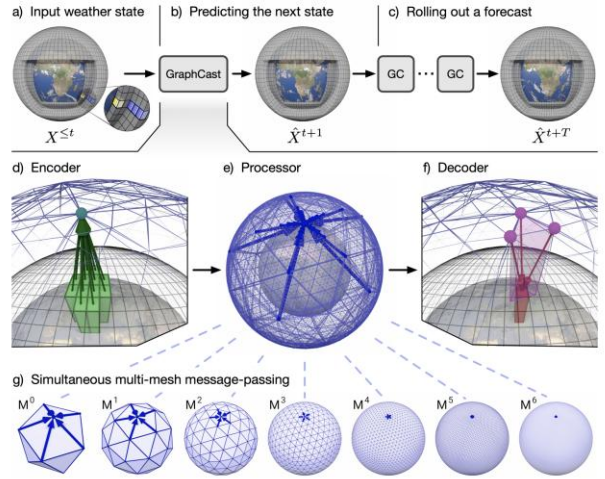
A PREPRINT

Jaideep Pathak, NVIDIA Corporation, Santa Clara, CA 95051
Shashank Subramanian, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
Peter Harrington, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
Sanjeev Raja, University of Michigan, Ann Arbor, MI 48109
Ashesh Chattopadhyay, Rice University, Houston, TX 77005
Morteza Mardani, NVIDIA Corporation, Santa Clara, CA 95051
Thorsten Kurth, NVIDIA Corporation, Santa Clara, CA 95051
David Hall, NVIDIA Corporation, Santa Clara, CA 95051
Zongyi Li, California Institute of Technology, Pasadena, CA 91125; NVIDIA Corporation, Santa Clara, CA 95051
Kamyar Azizzadenesheli, Purdue University, West Lafayette, IN 47907
Pedram Hassanzadeh, Rice University, Houston, TX 77005
Karthik Kashinath, NVIDIA Corporation, Santa Clara, CA 95051
Animashree Anandkumar, California Institute of Technology, Pasadena, CA 91125; NVIDIA Corporation, Santa Clara, CA 95051



Pangu-Weather: A 3D High-Resolution System for Fast and Accurate Global Weather Forecast

Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian✉, *Fellow, IEEE*



GraphCast: Learning skillful medium-range global weather forecasting

Remi Lam*,1, Alvaro Sanchez-Gonzalez*,1, Matthew Willson*,1, Peter Wirnsberger*,1, Meire Fortunato*,1, Alexander Pritzel*,1, Suman Ravuri[1], Timo Ewalds[1], Ferran Alet[1], Zach Eaton-Rosen[1], Weihua Hu[1], Alexander Merose[2], Stephan Hoyer[2], George Holland[1], Jacklynn Stott[1], Oriol Vinyals[1], Shakir Mohamed[1] and Peter Battaglia[1]
*equal contribution, [1]DeepMind, [2]Google

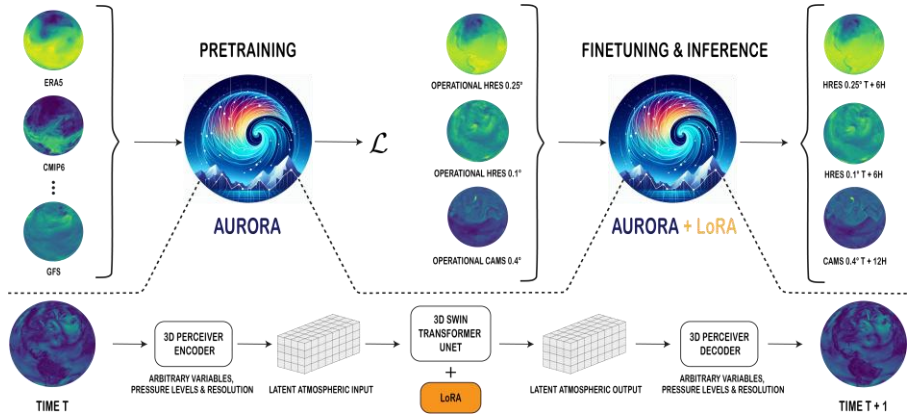
Aurora Atmospheric Foundation Model

AURORA: A FOUNDATION MODEL OF THE ATMOSPHERE

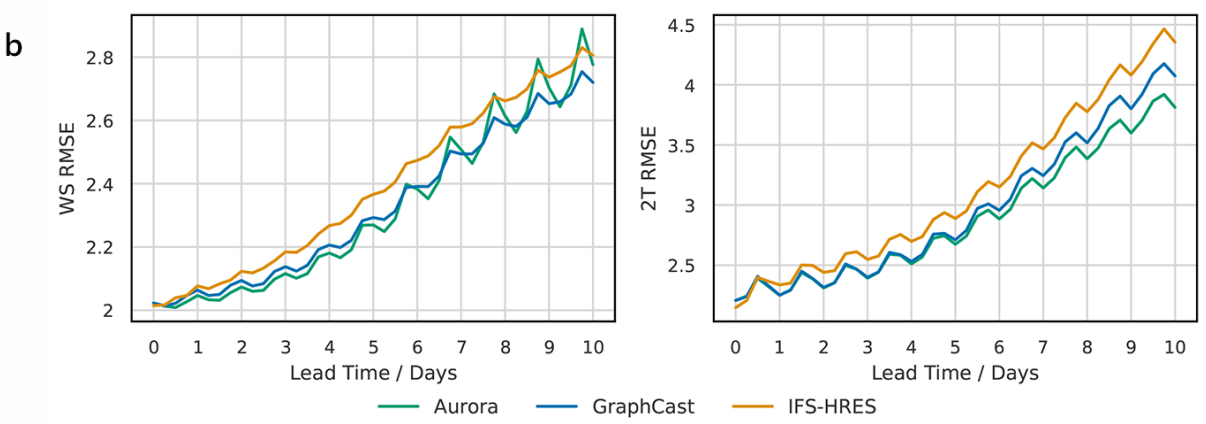
Cristian Bodnar[*,1], Wessel P. Bruinsma[*,1], Ana Lucic[*,1], Megan Stanley[*,1], Johannes Brandstetter[3,†], Patrick Garvan[1], Maik Riechert[1], Jonathan Weyn[2], Haiyu Dong[2], Anna Vaughan[4], Jayesh K. Gupta[5,†], Kit Tambiratnam[2], Alex Archibald[4], Elizabeth Heider[1], Max Welling[6,†], Richard E. Turner[1,4], and Paris Perdikaris[1]

[1]Microsoft Research AI for Science
[2]Microsoft Corporation [3]JKU Linz [4]University of Cambridge [5]Poly Corporation [6]University of Amsterdam

[*]Equal contribution [†]Work done while at Microsoft Research

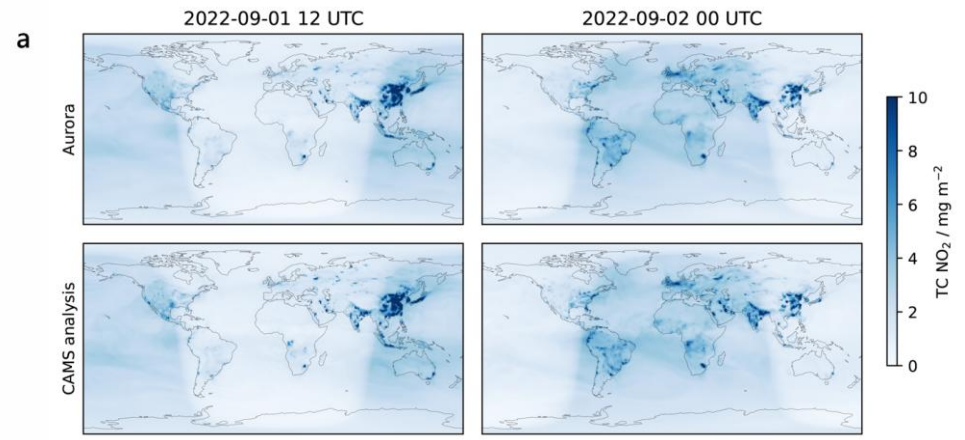


- Aurora Foundation model:
  - Can absorb/encode very different datasets (>1M hrs data)
  - Fast: 5000x speedup relative to IFS numerical forecasts
  - Can be applied/finetuned to many different tasks
    - Wind speeds, temperature,…
    - Air polution, concentration greenhouse gases,…



Windspeed Error

Surface Temperature Error

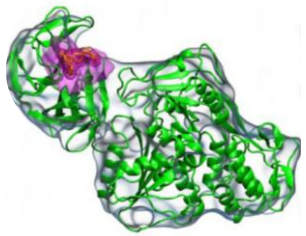


NO2 Air Polution

Molecules: From Drug Discovery to Carbon Capture

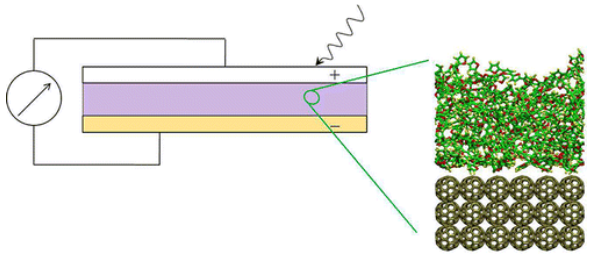
Everything material is made of molecules*

Molecules are at the root of solving many of the health, environmental and climate challenges we are facing today.



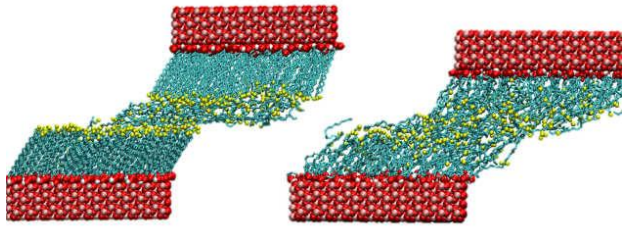
Drug discovery

Markus Reiher et al. PNAS 2017;114:29:7555-7560



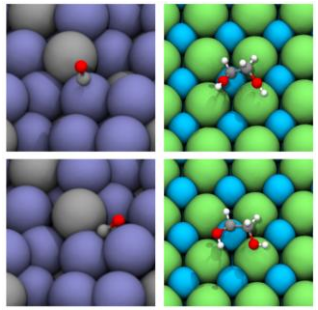
Photovoltaics

S.Y Reddy et al. Synthetic Metals 162, 23, 2012, 2117-2124



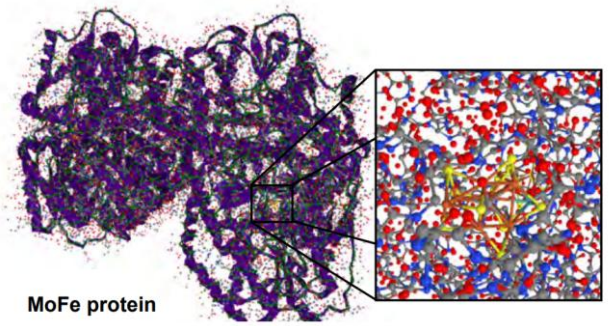
Tribology and lubricants

James Ewen, Tribology Group, Imperial College London



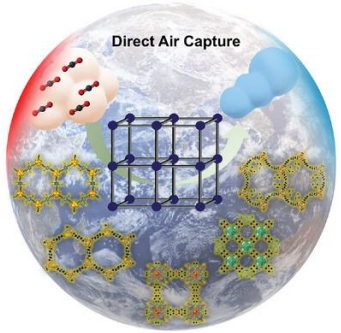
Catalyst design (e.g., fuel cells)

Lowik Chanussot et al. ACS Catal. 2021, 11, 10, 6059–6072



Nitrogen fixation

Shaher Bano Mirza et al. Journal of Molecular Graphics and Modelling 2016

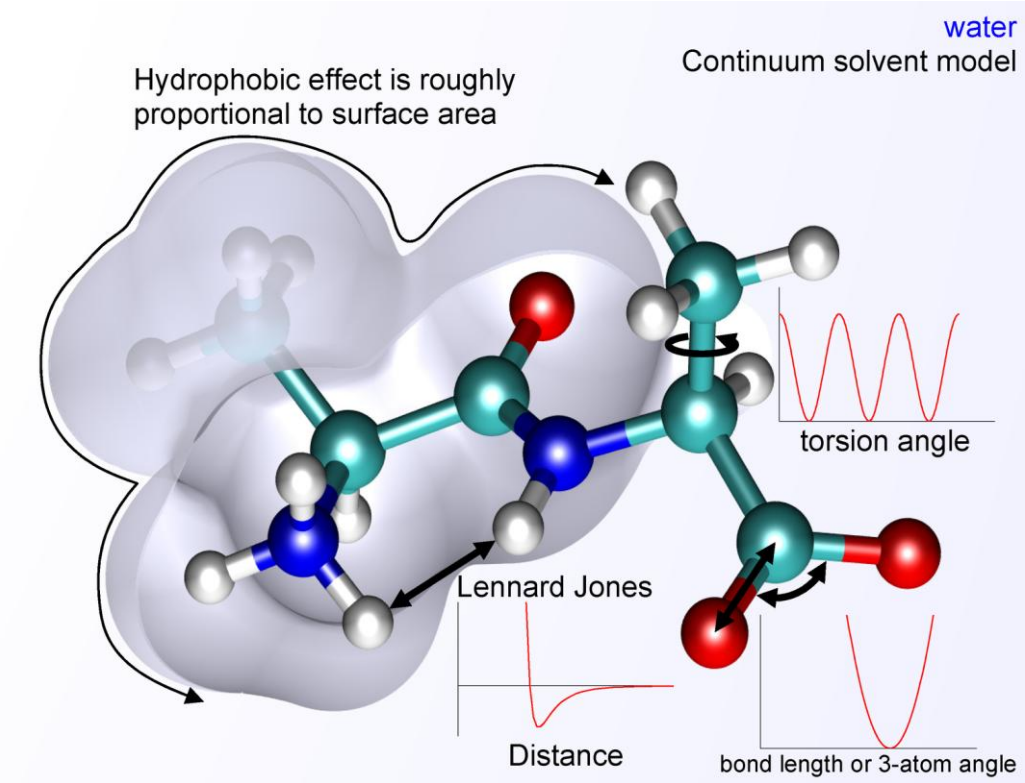


Carbon Capture

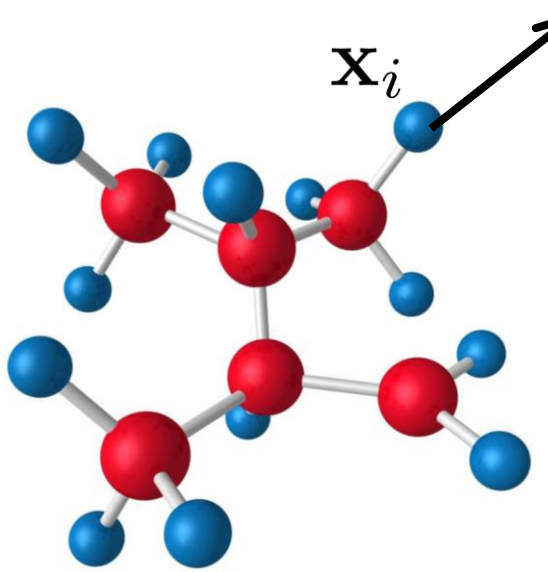
H Furukawa, KE Cordova, M O'Keeffe, OM Yaghi, Science 341 (6149)

* Except 4 fundamental forces (electromagnetic force, gravity and strong & weak nuclear forces), or unless you break them up (plasma, quarks/leptons)

Simulating and Designing Molecules: Three Challenges



$$\frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i$$
$$m_i \frac{d\mathbf{v}_i}{dt} = -\nabla_i U(\mathbf{x}_1, \ldots, \mathbf{x}_N)$$



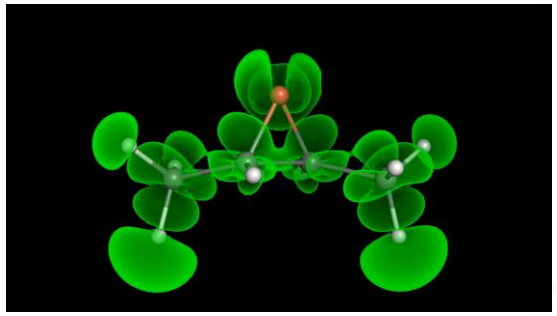
1) Forces are largely quantum mechanical

2) Molecular dynamics is chaotic

3) Search space of molecules and materials is enormous

How can AI help the Sciences?

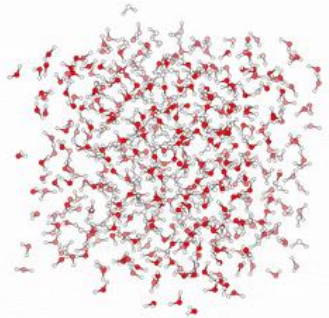
Schrodinger eqn.



$$i\hbar\frac{d}{dt}\psi(t) = H(t)\psi(t)\,.$$

(DFT+ML,
QVMC: Fermi & Pauli Net)

Langevin eqn.

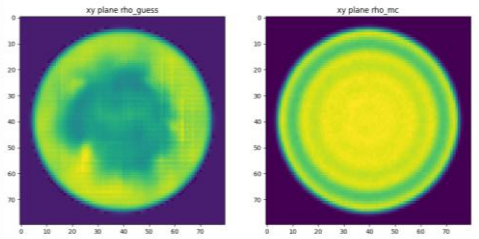


$$\dot{x} = v$$

$$\dot{v} = -\gamma v - \frac{1}{m}\frac{dU}{dx} + \sqrt{2B}\xi(t)$$

(MD+ML=ML Force Fields)

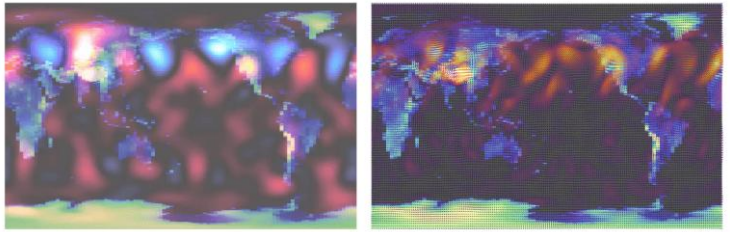
Fokker-Planck eqn.



$$\partial_t \rho(x,t) = \gamma \nabla_x \cdot \left(\rho(x,t)\nabla_x \frac{\delta\Omega[\rho]}{\delta\rho(x,t)}\right)$$

(Classical DFT + ML)

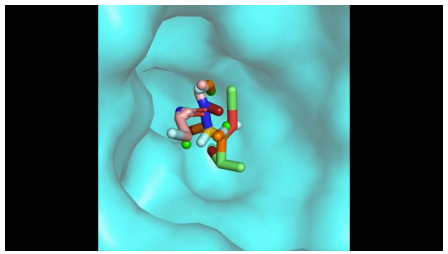
Navier Stokes eqn.



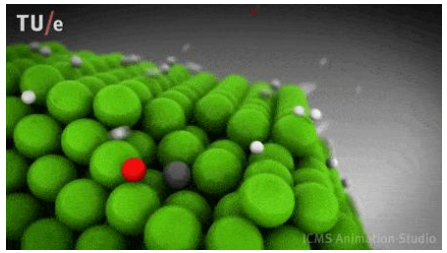
$$\partial \boldsymbol{u}/\partial t = \nu\,\Delta \boldsymbol{u} - (\boldsymbol{u}\cdot\boldsymbol{\nabla})\,\boldsymbol{u} + \boldsymbol{\nabla} p,$$
$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0,$$

(PDE Surrogates:
FourcastNet, GraphcastNet, Aurora)

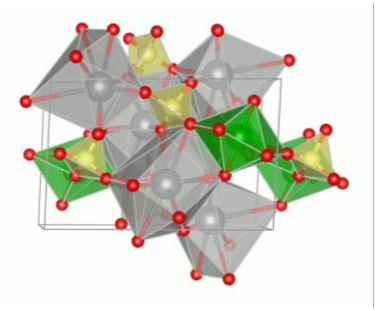
Drug Discovery:
Generating Molecules



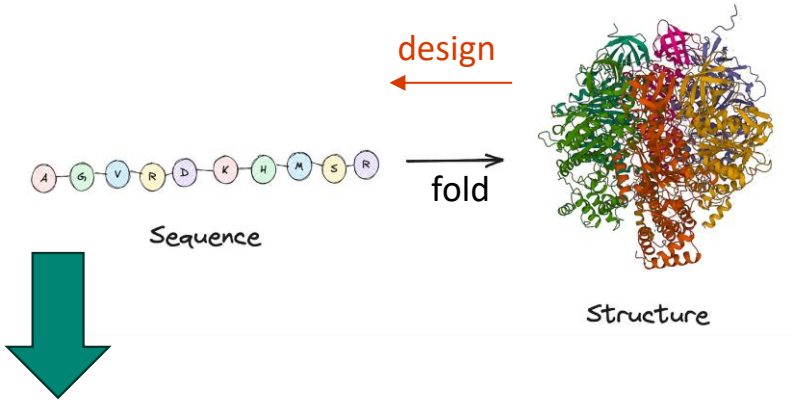
Catalysis:
Accelerate reactions.



Materials Science:
Generating materials

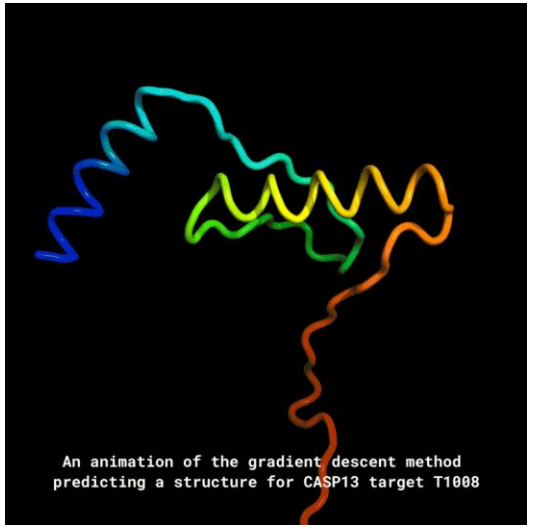


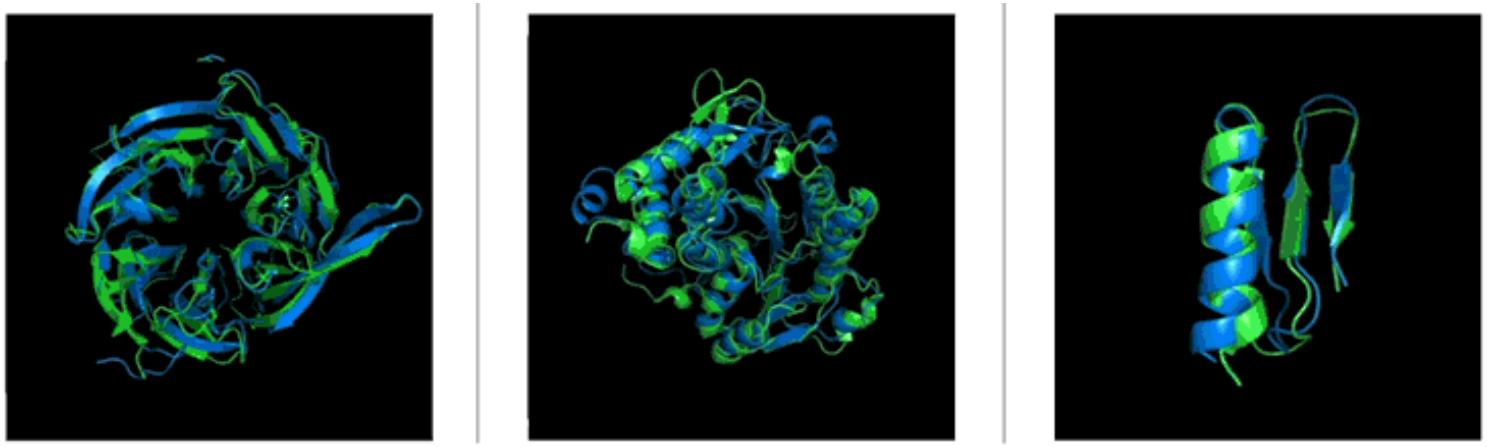
Google Deepmind's Alphafold





Nobel Prize Chemistry 2024

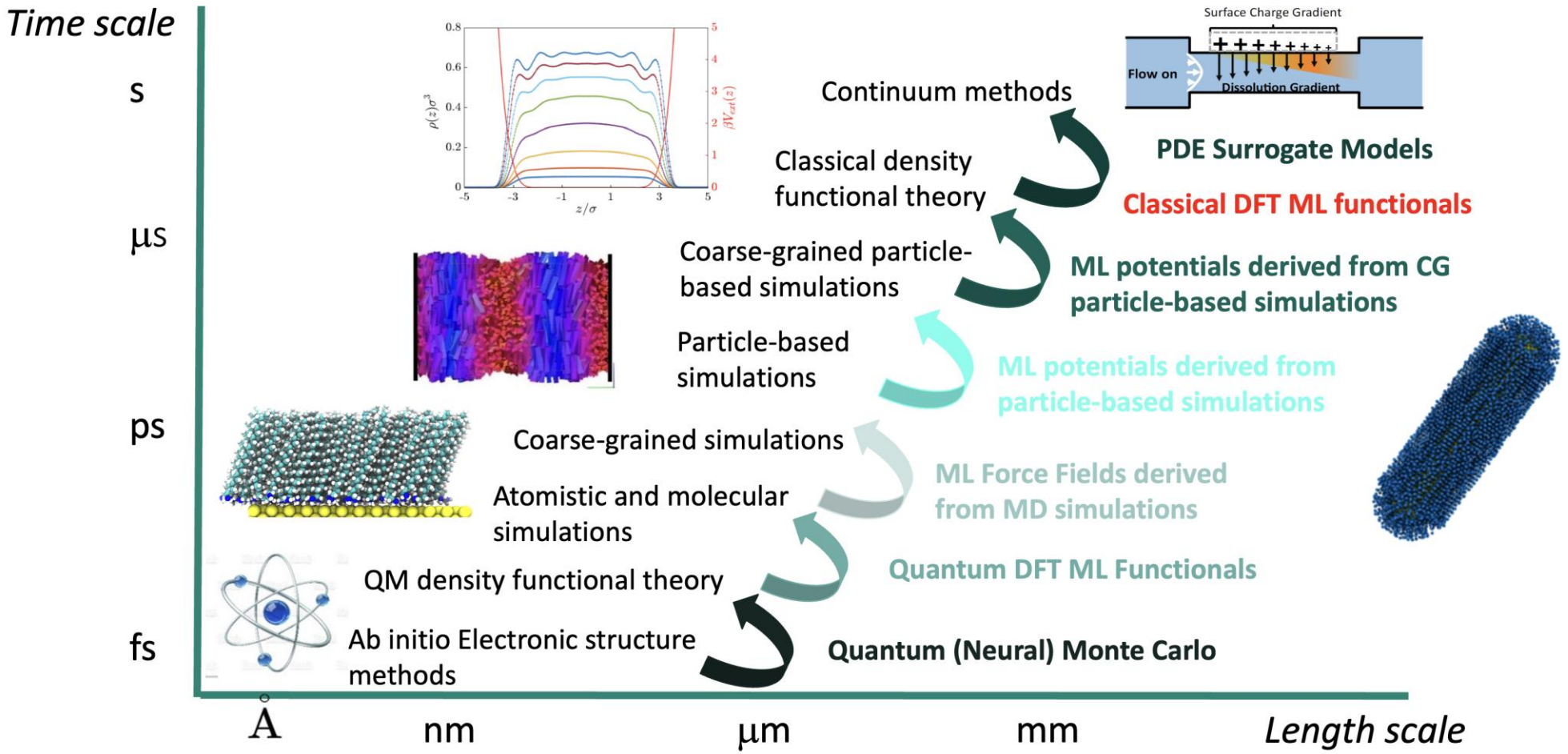




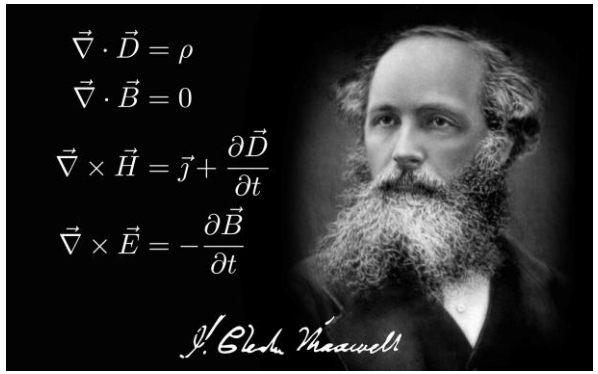
Clear evidence that AI is making an impact in the Sciences!

Potential ML Impact on Physics/Chemistry on all Scales

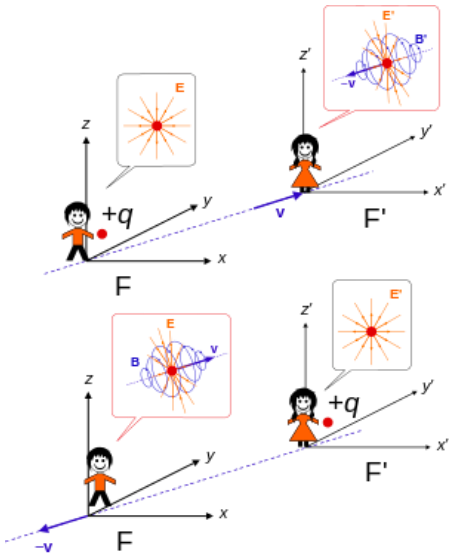
Slide credit: Marjolein Dijkstra



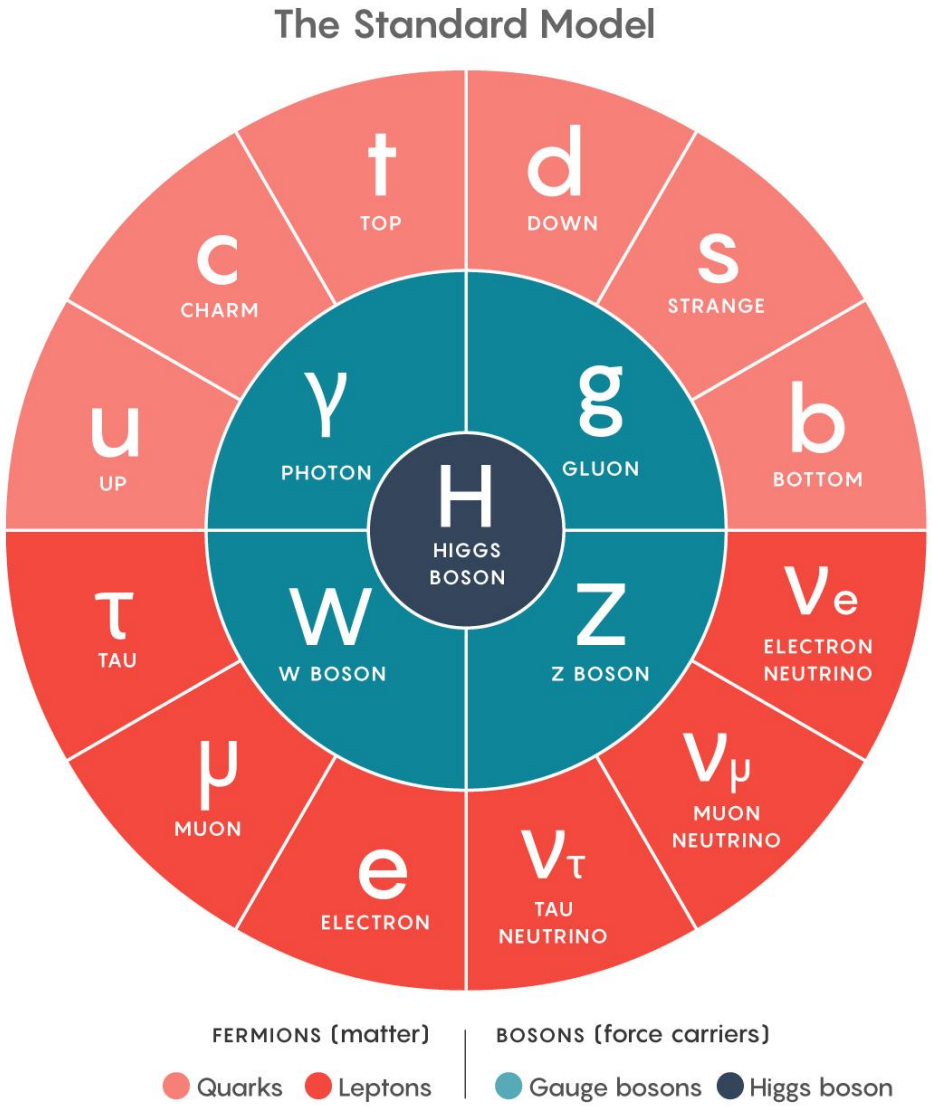
Symmetries & Equivariance in Physics

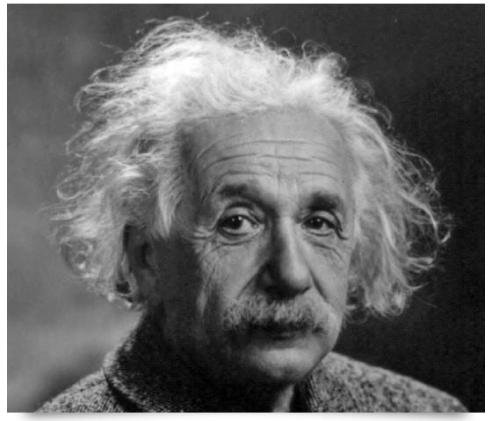


Electricity = Magnetism



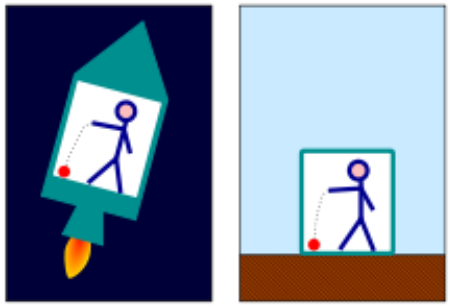
- ield.
-
-





Gravity = Acceleration

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} + \Lambda\, g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$



Graph Neural Network



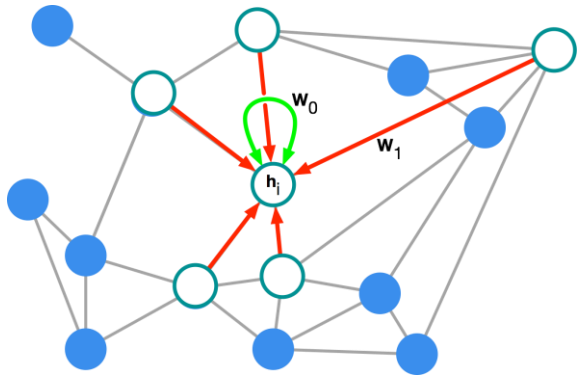
Thomas Kipf



Victor G. Satorras



Emiel Hoogeboom



GNN

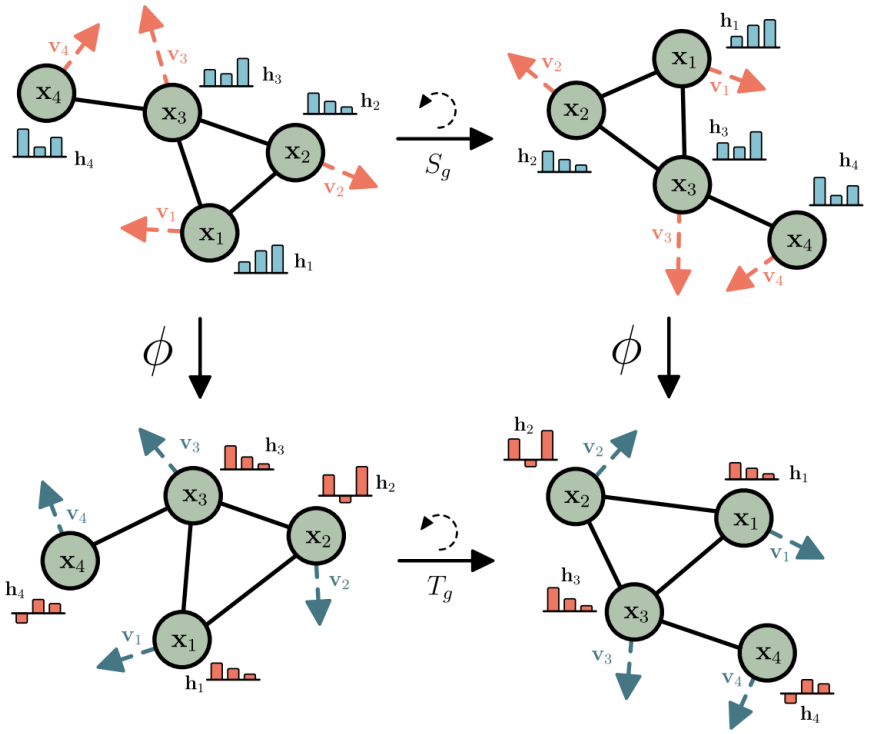


Figure 1. Example of rotation equivariance on a graph with a graph neural network ϕ

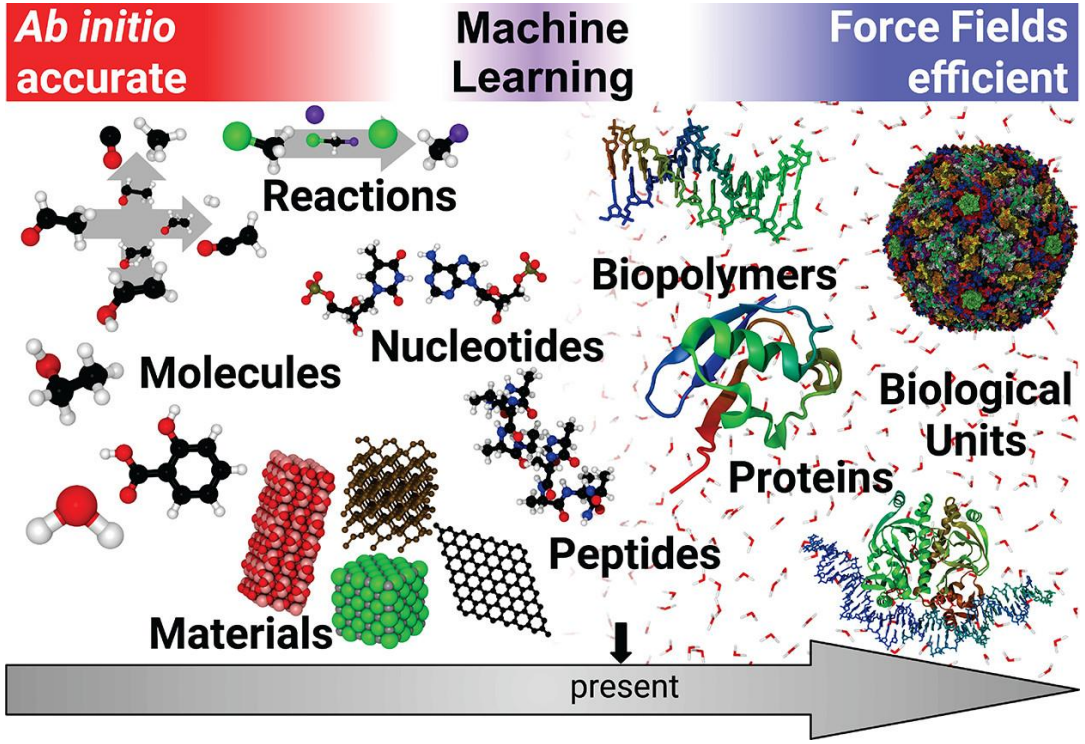
Equivariant GNN

Machine Learning (Equivariant) Force Fields

REVIEW | March 11, 2021

Machine Learning Force Fields

Oliver T. Unke, Stefan Chmiela, Huziel E. Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T. Schütt, Alexandre Tkatchenko*, and Klaus-Robert Müller*



SE(3)-Transformers: 3D Roto-Translation Equivariant Attention Networks

Fabian B. Fuchs*†
Bosch Center for Artificial Intelligence
A2I Lab, Oxford University
fabian@robots.ox.ac.uk

Daniel E. Worrall*
Amsterdam Machine Learning Lab, Philips Lab
University of Amsterdam
d.e.worrall@uva.nl

Volker Fischer
Bosch Center for Artificial Intelligence
volker.fischer@de.bosch.com

Max Welling
Amsterdam Machine Learning Lab
University of Amsterdam
m.welling@uva.nl

Steerable Equivariant Message Passing on Molecular Graphs

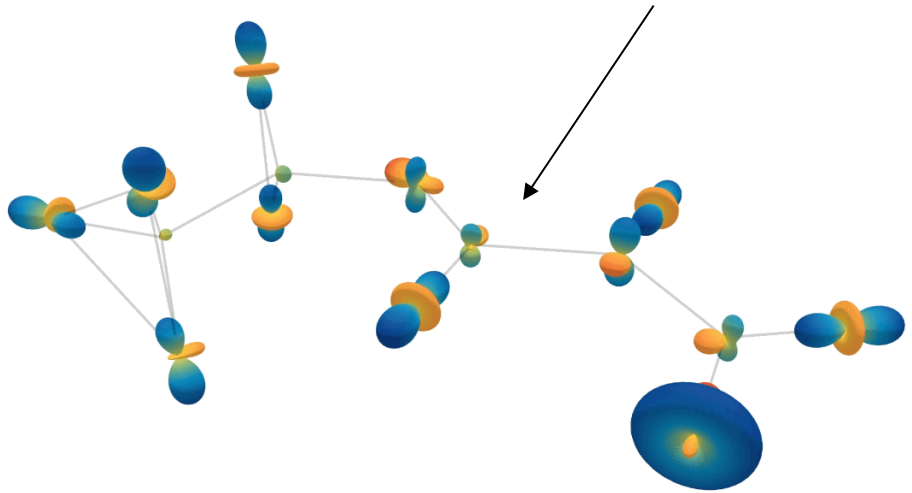
Johannes Brandstetter*
University of Amsterdam
Johannes Kepler University Linz
brandstetter@ml.jku.at

Rob Hesselink*
University of Amsterdam
r.d.hesselink@uva.nl

Elise van der Pol
University of Amsterdam
e.e.vanderpol@uva.nl

Erik Bekkers
University of Amsterdam
e.j.bekkers@uva.nl

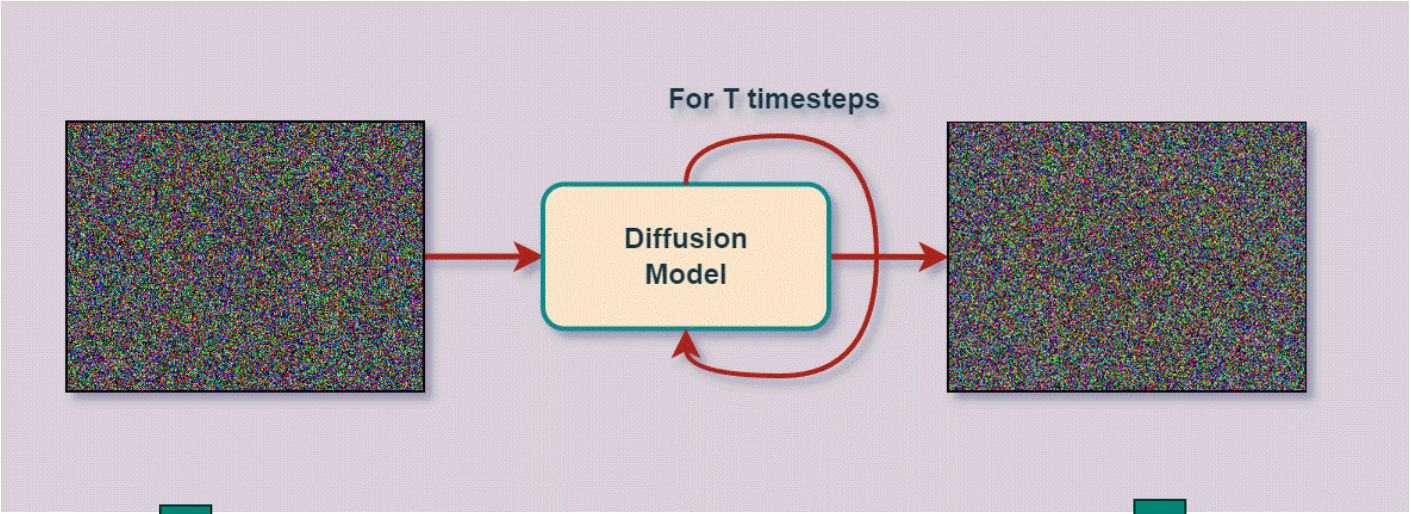
Max Welling
University of Amsterdam
Qualcomm AI Research
m.welling@uva.nl



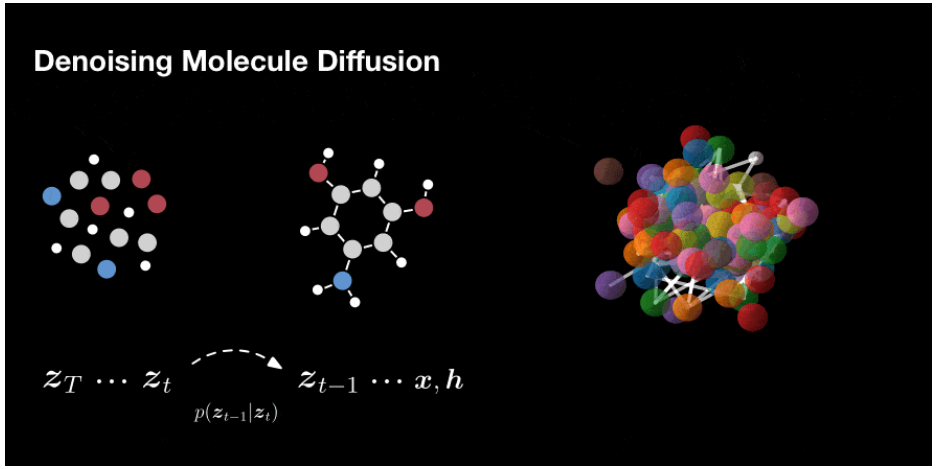
Fully SE(3) equivariant architecture

Spherical harmonics representations through the layers

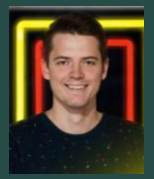
Generative AI



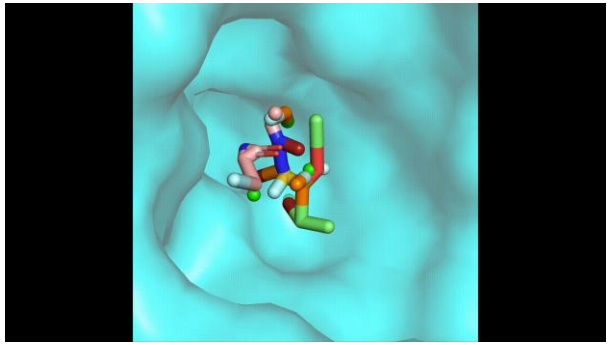




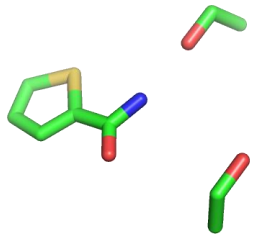
DiffLinker: Molecular Linker Design

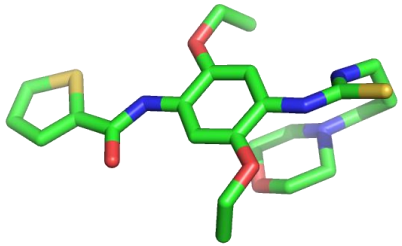
Arne Schneuing Ilia Igashov (+B. Correia, M. Bronstein & W. et al)



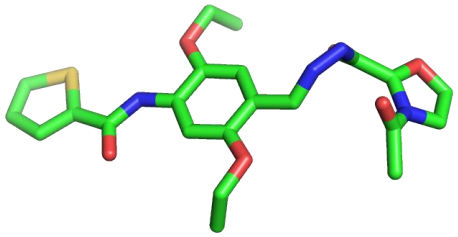
Input fragments

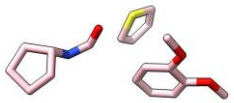
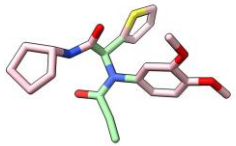
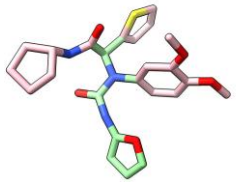
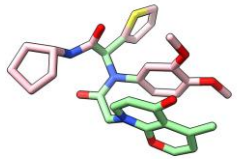
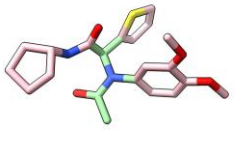
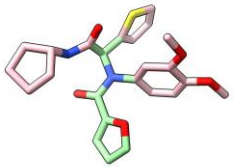
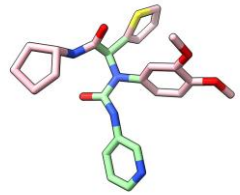
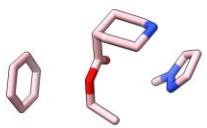
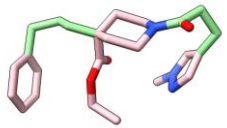
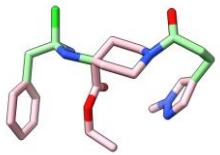
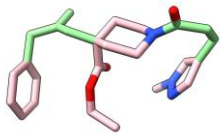
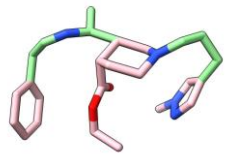
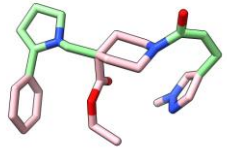
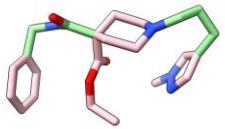


Reference molecule



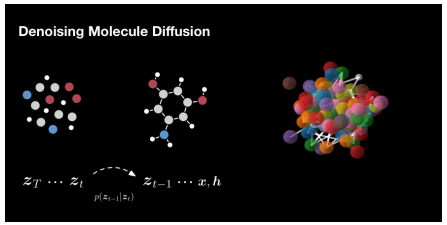
Generated molecule



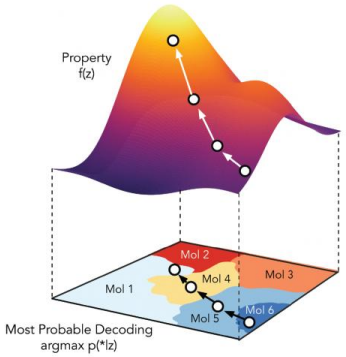
Input Fragments	True Molecule	DiffLinker Samples				
						
						

The Key Problem: Design = Search

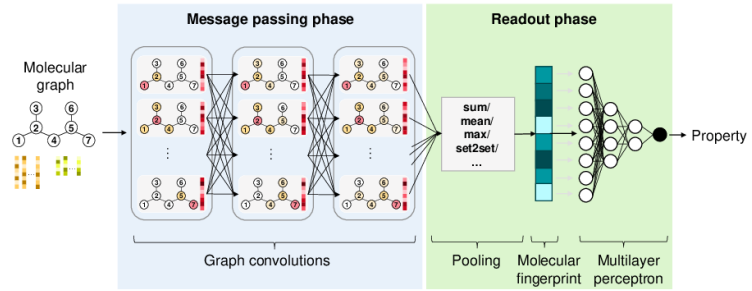
CuspAI



Generation new materials



Search through chemical space



Evaluate proprties

Retrain models

Store data



Database

Store data

Retrain models

A Golden Age of Materials

Time



Stone age



Bronze age



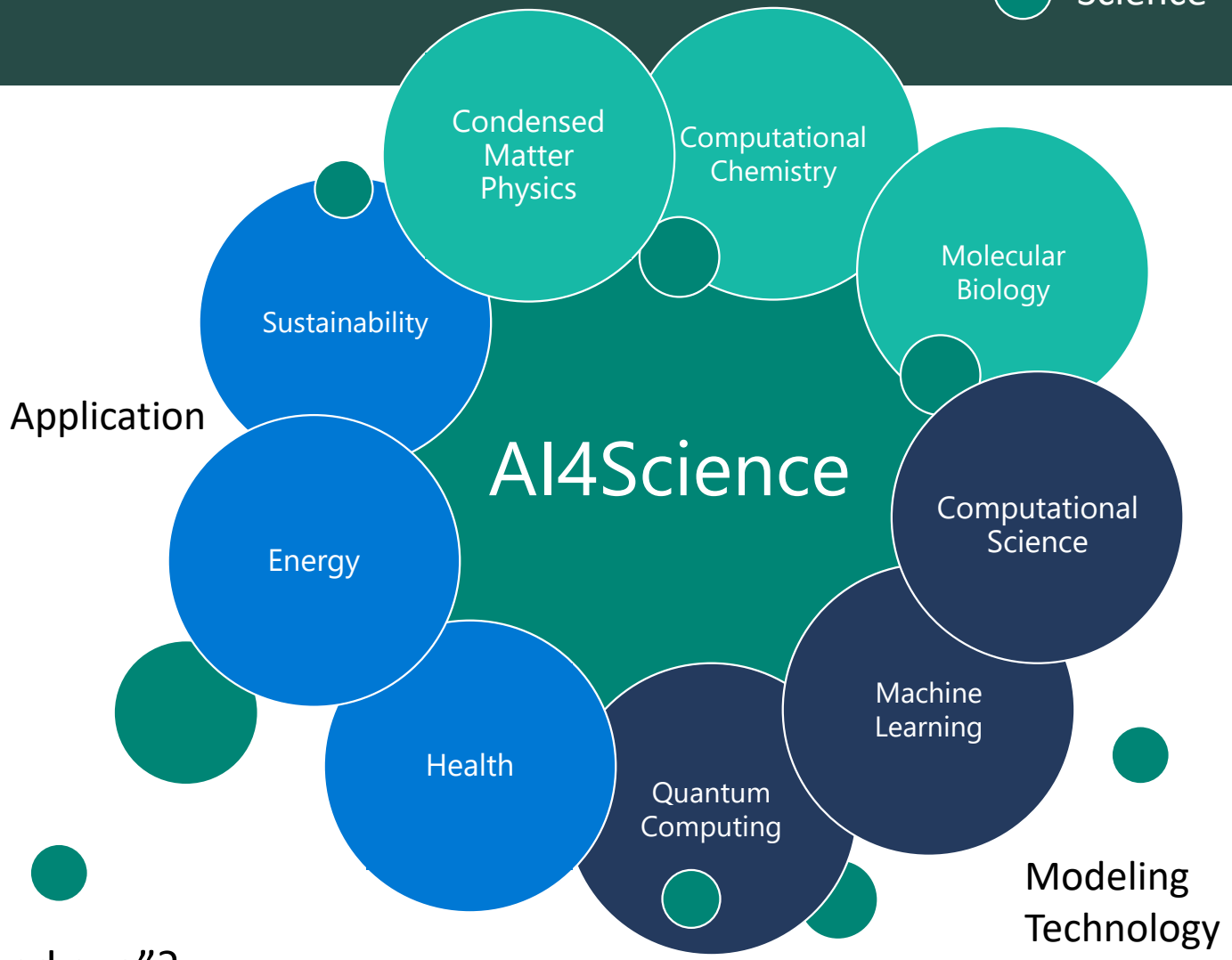
Iron age



Plastic age

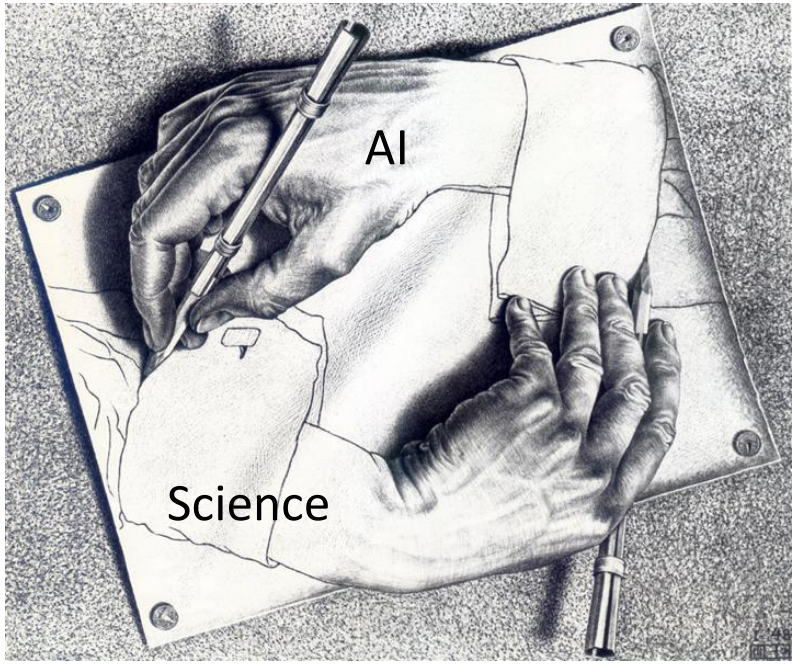


“Material-on-demand age”?



Conclusions

AI and Science are catalyzing each other.



ML is the new hammer for computational scientist.



Applications in health & sustainability: new drugs & materials.



(ChatGPT's impression of a carbon capture machine)